

# CAYLEY GRAPHS AND AUTOMATIC SEQUENCES

PIERRE GUILLOT

**ABSTRACT.** We study those automatic sequences which are produced by an automaton whose underlying graph is the Cayley graph of a finite group. For 2-automatic sequences, we find a characterization in terms of what we call *homogeneity*, and among homogeneous sequences, we single out those enjoying what we call *self-similarity*. It turns out that self-similar 2-automatic sequences (viewed up to a permutation of their alphabet) are in bijection with many interesting objects, for example dessins d'enfants (covers of the Riemann sphere with three points removed).

For any  $p$  we show that, in the case of an automatic sequence produced “by a Cayley graph”, the group and indeed the automaton can be recovered canonically from the sequence.

Further, we show that a rational fraction may be associated to any automatic sequence. To compute this fraction explicitly, knowledge of a certain graph is required. We prove that for the sequences studied in the first part, the graph is simply the Cayley graph that we start from, and so calculations are possible.

We give applications to the study of the frequencies of letters.

## 1. INTRODUCTION

**1.1. Basic definitions.** Let  $p \geq 2$  be an integer, which in practice will often be a prime. A  $p$ -*automaton* is, first and foremost, a directed graph on a finite set  $Q$ , whose elements are called the *states*; the following extra decoration is required :

- There is a distinguished state, called the *initial state*.
- The arrows are labeled with the integers  $i$  such that  $0 \leq i < p$ .
- The vertices are labeled using a map  $\tau: Q \rightarrow \Delta$ . Here  $\Delta$  is a finite set called the *alphabet*. Typically  $\tau$  will be surjective, but may very well fail to be injective.

Finally, the following property must be satisfied: out of each state (vertex), there is precisely one arrow labeled  $i$ , for each  $0 \leq i < p$ . For example, here is a 2-automaton.

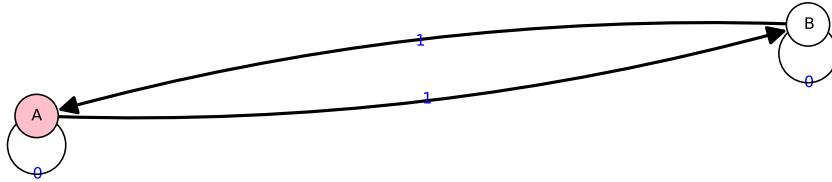


Fig. 1

In this example the alphabet is  $\Delta = \{A, B\}$ ; the initial state is the state bearing the label  $A$ . On our pictures we usually depict the initial state in a darker colour.

This definition is equivalent to others in the literature, and it will serve our purposes well, at the cost of (mildly) surprising some readers.

An automaton produces a sequence  $(a_n)_{n \geq 1}$  of elements of  $\Delta$  using the following recipe. Read the digits of  $n$  written in base  $p$ , from right to left, and follow the corresponding arrows in the graph, starting from the initial state ; the label of the state reached in this fashion is  $a_n$ . Using the automaton from Figure 1, one obtains the celebrated Thue-Morse sequence, starting with

$$BBABAABBAABABBABAABBAABBABAABB \dots$$

Many readers familiar with the Thue-Morse sequence will complain that there is an  $A$  missing at the beginning. However, we chose to define the sequence associated to an automaton for  $n \geq 1$  (rather than  $n \geq 0$ ), and we make no apology for this unorthodox decision: the results of the present paper will irremediably fail to hold for sequences defined from 0.

A sequence of elements of  $\Delta$  which is the output of at least one  $p$ -automaton is called *p-automatic*. There is a vast literature on automatic sequences (see [AS03]), and most results do not depend on whether the sequences start from 0 or 1.

The automaton on Figure 2 appeared in [RY15] and [Row15].

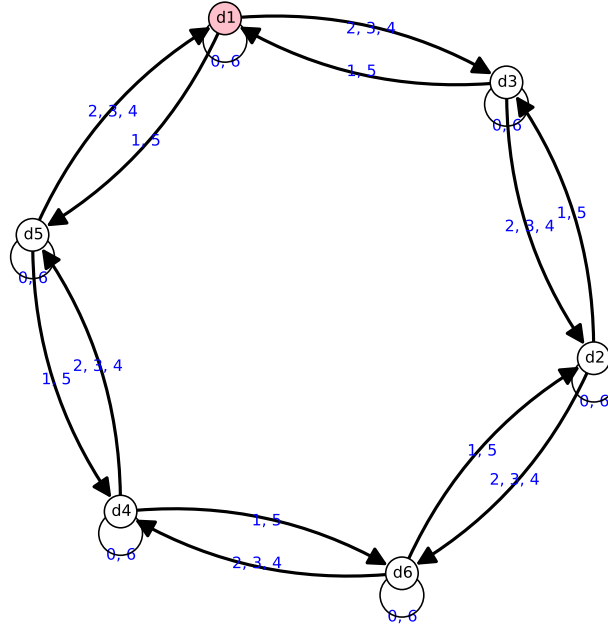


Fig. 2

On this picture we have written several labels next to a given arrow as a way of saving space. There are in fact, for example, three arrows going from  $d_1$  to  $d_3$ , with labels 2, 3 and 4 respectively.

The sequence produced by this automaton is  $(A(n) \bmod 7)_{n \geq 1}$  where  $A(n)$  is the Apéry number

$$A(n) = \sum_{k=0}^n \binom{n}{k}^2 \binom{n+k}{k}^2.$$

(More precisely if the output for a given  $n$  is  $d_i$ , then  $A(n) \bmod 7$  is  $i$ . We chose the alphabet to be  $\Delta = \{d_1, \dots, d_6\}$  rather than  $\{1, \dots, 6\}$  to emphasize that  $d_i$  is a formal symbol and that no arithmetic is performed with these outputs in the sequel.)

In [Row15], Rowland comments that this automaton is “particularly symmetric”. In this paper we investigate those sequences which are produced by automata with a lot of symmetry.

One may think of an automaton as a particular kind of computing machine, which outputs  $a_n$  when fed  $n$ . If one adopts this point of view, the question of symmetry is not natural: one hardly cares for symmetric computers or computer programs. However, it turns out that the *answer* to this surprising question is very simple and satisfying (especially for  $p = 2$ ).

First, we need to define what we mean by symmetry. Suppose  $G$  is a group with distinguished generators  $t_0, t_1, \dots, t_{p-1}$ . Its *Cayley graph* is built as follows: the set of vertices is  $G$  itself, and there is one arrow with the label  $i$  placed between  $g$  and  $gt_i$  (for all  $g \in G$  and  $0 \leq i < p$ ).

Both our examples above involve Cayley graphs. If we take the symmetric group  $S_2$ , generated by  $t_0 = I$  (the identity permutation) and  $t_1 = (1, 2)$ , we get precisely the graph underlying Figure 1. As for Figure 2, consider the permutations

$$t_0 = t_6 = I, \quad t_1 = t_5 = (1, 5, 4, 6, 2, 3), \quad t_2 = t_3 = t_4 = (1, 3, 2, 6, 4, 5),$$

and the group  $G$  that they generate (which is cyclic of order 6); what we have on the figure is precisely the corresponding Cayley graph.

Note that a Cayley graph always has a distinguished vertex, namely the identity of  $G$ . Also, the characteristic property for automata is satisfied *as well as it “dual”*, namely: at each vertex there is exactly one arrow with the label  $i$  going out, and also exactly one arrow with the label  $i$  coming in. However to turn a Cayley graph into an automaton, one needs to provide a map  $\tau: G \rightarrow \Delta$ .

**1.2. Two types of sequences with symmetry.** We propose to give necessary and sufficient conditions for an automatic sequence to be produced by an automaton whose underlying graph is a Cayley graph. These have a vertex-transitive group of automorphisms, and are thus very symmetric objects indeed.

Here is some notation to formulate these conditions. Let the integer  $p$  be fixed throughout. For  $i, j \geq 0$  with  $j < p^i$ , we define the subsequence  $a^{(i,j)}$  of the sequence  $a$  by

$$a_n^{(i,j)} = a_{p^i n + j}$$

for  $n \geq 1$  (not  $n \geq 0$ !). We let

$$N(a) = \{a^{(i,j)} : i, j \geq 0, j < p^i\}$$

be the set of such sequences. (As we shall recall below, it is well-known that  $N(a)$  is finite if and only if  $a$  is  $p$ -automatic.)

Further, a *relation* for the sequence  $a$  is a pair  $(i, j)$  such that  $a^{(i,j)} = a$ ; we exclude  $(i, j) = (0, 0)$ , which we do not consider as a relation. We write  $\text{rel}(a)$  for the set of all relations of  $a$ .

We will say that  $a$  is *homogeneous* if

$$\forall u \in N(a), \quad N(u) = N(a) \quad \text{and} \quad \text{rel}(u) = \text{rel}(a).$$

Further, call two sequences  $a$  and  $b$  *equivalent*, and write  $a \sim b$ , when there is a bijection  $\varphi: \Delta \rightarrow \Delta'$ , where  $\Delta$  resp.  $\Delta'$  is the alphabet of  $a$  resp.  $b$ , such that  $b_n = \varphi(a_n)$  for all  $n \geq 1$ . We call a sequence  $a$  *self-similar* if  $a \sim u$  for all  $u \in N(a)$ . Self-similar sequences are homogeneous; to see this, note that when  $u \in N(a)$  one always has  $N(u) \subset N(a)$ , and if  $a \sim u$  we draw  $|N(u)| = |N(a)|$  so that  $N(u) = N(a)$ , and  $\text{rel}(u) = \text{rel}(a)$  is then obvious.

For example, if  $a$  is the Thue-Morse sequence, then  $N(a)$  has just two elements, namely  $a^{0,0} = a$  and  $a^{1,1} = (a_{2n+1})_{n \geq 1}$ . We have  $a^{1,0} = (a_{2n})_{n \geq 1} = a$  so  $(1, 0) \in$

$\text{rel}(a)$ , giving an example of relation; more importantly, we have  $a^{1,1} \sim a$  since  $a^{1,1}$  is simply obtained by switching  $A$  and  $B$ . So  $a$  is self-similar.

Our first result is the following (results for general  $p$ -sequences are presented below).

**THEOREM 1.1** – *Let  $(a)$  be a 2-automatic sequence. Then  $(a)$  is homogeneous if and only if it can be produced by an automaton whose underlying graph is the Cayley graph associated to a group  $G$ , and such that the map  $\tau: G \rightarrow \Delta$  has the following property: the subgroup  $H$  of elements  $h \in G$  verifying  $\tau(hg) = \tau(g)$ , for all  $g \in G$ , is normal in  $G$ .*

*Moreover,  $(a)$  is self-similar if and only if it can be produced by an automaton whose underlying graph is the Cayley graph associated to a group  $G$ , and such that the alphabet can be identified with  $G/K$  for some subgroup  $K$ , with  $\tau: G \rightarrow G/K$  the natural map.*

(In the text there is also an even weaker type of sequence, corresponding to a weaker type of symmetry. The automata involved are those which have been called *permutation automata* or *reversible automata* in the literature.)

The second statement is probably the most satisfying: a kind of symmetry in the graph translates precisely into another sort of symmetry within the sequence. Note also that the condition on the subgroup  $H$ , in the first statement, is satisfied if  $\tau$  is conjugation-invariant in the sense that  $\tau(y^{-1}xy) = \tau(x)$ . For example  $\tau$  may be the character of a representation. Finally, we point out that one may replace  $G$  by  $G/H$  (since  $H$  is normal), and obtain a smaller automaton still producing  $a$ . We turn to questions of minimality.

**1.3. Canonical automata.** These results are entangled with another question. When  $a$  is an automatic sequence, is there a way to *canonically* construct an automaton which produces  $a$ ? It turns out that the answer is affirmative for homogeneous 2-sequences (and so also self-similar sequences).

We shall in fact define an oriented graph  $\Gamma(a)$  from any  $p$ -automatic sequence  $a$  (for any  $p$ ), whose set of vertices is  $N(a)$ . We also define a monoid  $G(a)$  of self-maps of  $N(a)$ , having generators  $t_0, \dots, t_{p-1}$ , and  $\Gamma(a)$  is the “Schreier graph” of this monoid (we define Schreier graphs below). There is a distinguished vertex in  $\Gamma(a)$ , namely  $a$  itself. We shall prove that whenever  $G(a)$  is actually a group, and not just a monoid, then there is a natural map  $\tau: N(a) \rightarrow \Delta$  (indeed  $\tau(u)$  is the first term of the sequence  $t_1^{-1}(u)$ ). Thus  $\Gamma(a)$  is a full-blown automaton in this case.

We establish the following.

**THEOREM 1.2** – *Let  $(a)$  be a 2-automatic sequence. Then  $(a)$  is homogeneous if and only if  $G(a)$  is a group and  $\Gamma(a)$  is its Cayley graph. In this case  $\Gamma(a)$  produces  $(a)$ .*

For a general  $p$ , the statements of the two theorems stated so far do not hold in such generality. Indeed consider the sequence  $a_n =$  the leftmost digit of  $n$  when written in base  $p$ . One can show that  $a$  is  $p$ -automatic, with  $N(a) = \{a\}$ , so that the graph  $\Gamma(a)$  has only one vertex. Clearly it cannot be turned into an automaton producing  $a$ , unless  $p = 2$ .

However, a great deal remains true. Here, let us state the following.

**THEOREM 1.3** – *Let  $G$  be a group with distinguished generators  $t_0, \dots, t_{p-1}$ . Suppose the corresponding Cayley graph is turned into an automaton, with initial state 1, by means of the map  $\tau: G \rightarrow G/K$  for some subgroup  $K$ . Let  $(a)$  be the  $p$ -automatic sequence produced.*

*Then  $(a)$  is self-similar, the monoid  $G(a)$  is a group and  $\Gamma(a)$  is its Cayley graph. If moreover the intersection of all the conjugates of  $K$  is trivial, then  $G$  can*

be identified with  $G(a)$ , and  $\Gamma(a)$  can be identified with the graph underlying the automaton.

A natural choice is  $K = \{1\}$ , and the alphabet is then  $G$  itself. Both examples at the beginning of this introduction are of this kind: we have noticed that the graphs on Figure 1 and Figure 2 are Cayley graphs, and since the labeling map  $\tau$  is injective, we may see it as simply giving names to the elements of the group (for example on Figure 2 the unit of  $G$  is called  $d_1$ , the element  $t_1 = t_5$  is called  $d_5$ , the element  $t_2 = t_3 = t_4$  is called  $d_3$ ).

*Remark 1.4.* If we combine the results stated so far, for  $p = 2$ , we see that a 2-automatic, self-similar sequence determines, and is entirely determined up to equivalence by, a finite group  $G$ , two generators  $t_0$  and  $t_1$ , and a conjugacy class of subgroups such that the intersection of all the subgroups in the class is trivial. This is tantamount to specifying a conjugacy class of subgroups of finite index in the free groups on two generators  $\langle t_0, t_1 \rangle$ . In turn, many interesting objects, comprising the theory of *dessins d'enfants* as in [Gui14], are in bijection with 2-automatic, self-similar sequences.

**1.4. Rational fractions.** We give an application to the study of certain rational fractions associated to automatic sequences.

The following holds in full generality. It may be known to the experts, but the author was not able to find a statement in the literature.

**THEOREM 1.5** – *Let  $a$  be a  $p$ -automatic sequence, and assume that the elements of the alphabet  $\Delta$  are taken in a ring. Define*

$$L(a, x) = \sum_{n \geq 1} a_n x^{\ell(n)},$$

where  $\ell(n)$  is the length of  $n$  when written in base  $p$ . Then  $L(a, x)$  is a rational fraction.

Moreover, there is an explicit formula for computing  $L(a, x)$ , involving the incidence matrix of the graph  $\Gamma(a)$ .

The reason why this result has not received much attention is perhaps that the graph  $\Gamma(a)$ , in general, is difficult to determine explicitly. This is where our previous considerations will be useful: we have just given a recipe for constructing an automatic sequence from a group  $G$  with distinguished generators, in such a way that  $\Gamma(a)$  is just the Cayley graph of  $G$  (and thus is explicitly known to us from the outset). We have also pointed out that our two running examples are of this kind.

For the Thue-Morse sequence  $a$ , seeing the alphabet  $\{A, B\}$  as a subset of the ring  $\mathbb{Z}[A, B]$ , we find

$$L(a, x) = \frac{-Ax^2 + Bx^2 - Bx}{2x - 1}.$$

For the 7-automatic sequence  $b$  produced by the automaton on Figure 2, we work with the ring  $\mathbb{Z}[d_1, d_2, \dots, d_6]$  and find:

$$L(b, x) = \frac{d_1 P_1 + d_2 P_2 + d_3 P_3 + d_4 P_4 + d_5 P_5 + d_6 P_6}{-441x^6 - 336x^5 - 300x^4 + 128x^3 + 24x^2 - 12x + 1}$$

where

$$\begin{aligned} P_1 &= 497x^6 + 380x^5 + 136x^4 - 80x^3 + 2x^2 + x, \\ P_2 &= -112x^6 + 96x^5 + 70x^4 - 63x^3 + 9x^2, \\ P_3 &= 28x^6 - 148x^5 + 102x^4 + 42x^3 - 27x^2 + 3x, \\ P_4 &= 98x^6 - 179x^5 + 105x^4 - 28x^3 + 4x^2, \\ P_5 &= -203x^6 + 123x^5 + 68x^4 + 28x^3 - 18x^2 + 2x, \end{aligned}$$

and

$$P_6 = 70x^6 + 70x^5 - 175x^4 + 35x^3.$$

These rational fractions contain information, in particular, about the frequencies of letters. Write  $a[n, d_i]$  for the number of occurrences of the symbol  $d_i$  among the first  $n$  terms of the sequence  $a$ . The frequency of  $d_i$  is

$$\lim_{n \rightarrow \infty} \frac{a[n, d_i]}{n},$$

when this limit exists. We will be able to prove or disprove the existence of

$$\lim_{n \rightarrow \infty} \frac{a[p^n, d_i]}{p^n},$$

as well as evaluate its value, when  $a$  is  $p$ -automatic and we know  $L(a, x)$ . Let us be content, in this introduction, with the following.

**THEOREM 1.6** – *Suppose  $a$  is  $p$ -automatic, on the alphabet  $\{d_1, \dots, d_k\}$ , and let*

$$L(a, x) = \frac{\sum_i d_i P_i(x)}{D(x)}$$

where  $P_i(x), D \in \mathbb{Z}[x]$ . Assume that the roots of  $D$  have absolute value  $\geq \frac{1}{p}$ , that the only root of  $D$  of absolute value  $\frac{1}{p}$  is  $\frac{1}{p}$  itself, and that this root is simple. Then

$$\lim_{n \rightarrow \infty} \frac{a[p^n, d_i]}{p^n} = \frac{P_i(\frac{1}{p})}{\sum_j P_j(\frac{1}{p})}.$$

For our sequence  $b$  related to the Apéry numbers, the denominator is

$$(-1) \cdot (3x + 1) \cdot (7x - 1) \cdot (x^2 + x + 1) \cdot (21x^2 - 9x + 1),$$

and the theorem applies. The polynomials  $P_i$  might look very different from one another, but for all  $i$  we have

$$P_i\left(\frac{1}{7}\right) = \frac{570}{16807}.$$

Thus the “frequencies” all agree, and since they sum up to 1, they must be equal to  $\frac{1}{6}$ .

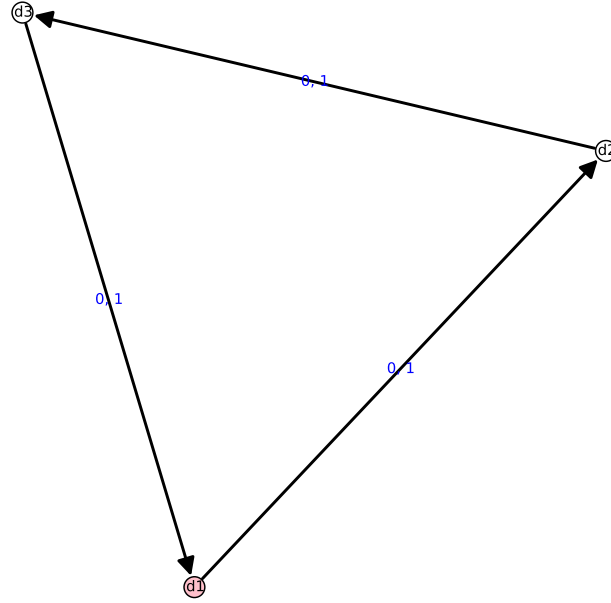
It is tempting to presume that, a self-similar sequence being so symmetric, the frequencies will always be  $\frac{1}{n}$  where  $n$  is the order of the group (= the size of the alphabet, the number of vertices). However we have a counter-example, of a 2-automatic sequence, which is self-similar with  $G(a)$  of order 8, for which the denominator has both  $\frac{1}{2}$  and  $-\frac{1}{2}$  as roots. Our more general statement implies that the limit (as above) does not exist, the values oscillating between  $\frac{1}{6}$  and  $\frac{1}{12}$ . Yet the average is still  $\frac{1}{2}(\frac{1}{6} + \frac{1}{12}) = \frac{1}{8}$ .

**1.5. Organization.** We complete the definitions in Section 2. The theorems stated in the introduction are proved in Section 3, albeit in a different order. Section 4 presents the rational fractions associated to automatic sequences, and their application to the computation of frequencies.

## 2. A FEW MORE PRELIMINARIES

The rather long introduction contained a number of definitions which will not be repeated. In this section we complete the set of definitions and make a few technical points.

**2.1. The need for sequences starting from 1.** Consider the following 2-automaton:



The usual practice is to associate to this automaton a sequence  $(a_n)_{n \geq 0}$  following the recipe we gave in the introduction for  $n \geq 1$ , and with  $a_0 =$  the initial state. We leave it as an exercise to the reader to check that  $a$  has then 9 subsequences of the form  $(a_{2^i n + j})_{n \geq 0}$  with  $j < 2^i$ ; moreover, these 9 sequences come in 3 groups, each consisting of three sequences which agree for  $n \geq 1$  but have different 0-th terms.

By contrast, if we consider  $(a_n)_{n \geq 1}$  and the subsequences belonging to  $N(a)$  as above, then there are just three of them. In general one has:

**LEMMA 2.1** – *Let  $Q$  be the set of states of a  $p$ -automaton producing the sequence  $(a_n)_{n \geq 1}$ . The map*

$$Q \longrightarrow N(a),$$

*which associates to  $q \in Q$  the sequence obtained by taking  $q$  to be the initial state, is surjective.*

*Proof.* Let  $q_0$  be the original initial state (producing  $a$ ). To produce the sequence  $(a_{p^i n + j})_{n \geq 1}$ , follow the arrows from  $q_0$  according to the digits of  $j$  written in base  $p$ , padded with 0's on the left so that  $i$  moves are made.  $\square$

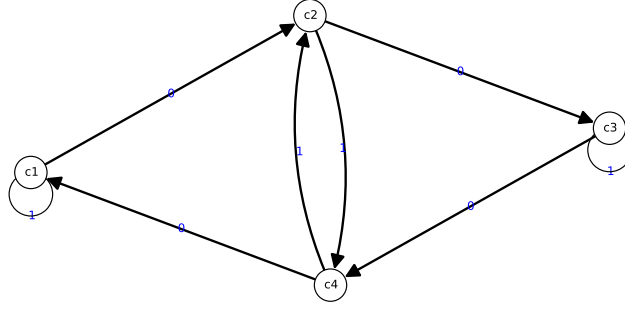
The example above shows that the lemma would fail for sequences starting from 0, since  $Q$  has 3 elements and we have found 9 subsequences.

On the other hand, for sequences starting from 1, the lemma proves that  $N(a)$  is finite when  $a$  is automatic. The converse is also true: when  $N(a)$  is finite, then  $a$  is  $p$ -automatic (note that the number  $p$  is used in the definition of  $N(a)$  even if it is absent from the notation). The classical proof (see for example [AS03]) is usually given for sequences starting from 0, but it is trivial to deduce the same statement for sequences starting from 1.

From now on, all sequences will start from 1. We may use the simple phrase “subsequence of  $a$ ” to mean specifically a subsequence of the form  $(a_{p^i n + j})_{n \geq 1}$  with  $j < p^i$ , when the context makes it clear.

**2.2. Schreier graphs.** We shall introduce sequences which are more general than homogeneous sequences (namely sequences “with global relations of all types”, see below), corresponding to a class of graphs more general than Cayley graphs.

Let  $G$  be a group (typically finite) with distinguished generators  $t_0, t_1, \dots, t_{p-1}$ . If  $G$  acts on a finite set  $Q$ , then one may form the *Schreier graph*, whose set of vertices is  $Q$ , and which includes a directed arrow bearing the label  $i$  between each  $q \in Q$  and its image under the action of  $t_i$ , for  $0 \leq i < p$ . For example, when  $Q = G$  and the action is simply multiplication (on the right), the corresponding Schreier graph is just the Cayley graph. On the other hand, if  $G$  is the dihedral group of orthogonal, planar transformations preserving a square, generated by a rotation  $t_0$  and a symmetry  $t_1$ , then the Schreier graph corresponding to the action on the four corners of the square is



It is easy to decide whether a given directed graph is a Schreier graph. This happens if and only if the property observed above for Cayley graph holds, namely, at each vertex there is exactly one arrow with the label  $i$  going out, and also exactly one arrow with the label  $i$  coming in. One can then construct a permutation  $t_i$  of  $Q$  whose action is dictated by the arrows in the graph, and define  $G$  to be the group generated by the  $t_i$ 's.

Let us say that a directed graph is *connected* if there is at least one vertex  $q_0$  such that any other vertex  $q$  can be reached from  $q_0$  by following directed arrows. The next lemma is well-known, and pretty trivial, but we include it for convenience.

**LEMMA 2.2** – *Let  $G$  be a finite group, with distinguished generators, acting on the set  $Q$ . Then the corresponding Schreier graph can be identified with the Cayley graph (non-canonically) if and only if it is connected and  $|Q| = |G|$ .*

*Proof.* Let  $q_0$  be as in the definition of connectedness (the arbitrary choice is why the identification will not be canonical). The map  $G \rightarrow Q$  which maps  $g$  to  $q_0^g$  (the result of letting  $g$  act on  $q_0$ ) is surjective by connectedness, and so also injective for reasons of cardinality. The desired identification follows.

The converse holds trivially. Moreover we see *a posteriori* that any  $q_0 \in Q$  could have been taken.  $\square$

Finally, note that the above definitions make sense if  $G$  is only a monoid rather than a group. The graph  $\Gamma(a)$ , alluded to in the introduction, is precisely the Schreier graph of a certain monoid. We turn to this.

**2.3. The graph  $\Gamma(a)$ .** Let  $a$  be  $p$ -automatic. We define maps

$$t_i: N(a) \longrightarrow N(a)$$

by

$$t_i(u) = (u_{pn+i})_{n \geq 1}$$



for  $0 \leq i < p$ . (The notation  $t_i$  is for “times  $p$  plus  $i$ ”.) The monoid generated by these (a submonoid of the monoid of all self-maps of  $N(a)$ ) will be written  $G(a)$ . We are particularly interested in situations when  $G(a)$  is a group, or equivalently when each  $t_i$  is a bijection, hence the notation. Typical elements of  $G(a)$  will be written  $g$  or  $h$ .

Many formulae will be simplified by the following convention. We define the composition of  $G(a)$  as follows:  $gh$  means first  $g$  and then  $h$ . Accordingly we will write  $u^g$  rather than  $g(u)$  (for  $g \in G(a)$ ,  $u \in N(a)$ ), and we have  $u^{g^h} = (u^g)^h$ .

Place an arrow with the label  $i$  between  $u$  and  $u^{t_i}$ , for all  $u$  and all  $i$ . We now have the directed graph  $\Gamma(a)$ , with labeled arrows, and also with the distinguished vertex  $a$ . Note that  $\Gamma(a)$  is the Schreier graph for the action of the monoid  $G(a)$  on the set  $N(a)$ .

It is important to realize that  $\Gamma(a)$  is connected. More precisely, there is a sequence of directed arrows leading from  $a$  to any  $u \in N(a)$ . This will follow from the computations which we describe now.

**2.4. Basic calculations.** The following calculations will be used many times, often implicitly. Let  $p$  be fixed, as ever. We define operations  $t_i$  for  $0 \leq i < p$  on polynomials by

$$f^{t_i} = f(pX + i) \quad \text{for } f \in \mathbb{Z}[X].$$

When  $w$  is a word in the alphabet  $\{t_0, \dots, t_i\}$ , then  $f^w$  has the obvious meaning. We are solely interested in the polynomials  $X^w$ , which are all of the form  $p^i X + j$  with  $j < p^i$ , by an immediate induction; we claim that each such polynomial is obtained for some  $w$ . This will prove that  $\Gamma(a)$  is connected.

In fact, we shall be more precise. To the polynomial  $p^i X + j$  with  $j < p^i$  we associate a word in the alphabet  $\{0, 1, \dots, p-1\}$ , denoted  $[p^i X + j]$ , by the following rule: write  $j$  in base  $p$ , and pad the results with 0's on the left so as to have  $i$  digits in total. For example with  $p = 2$  one has  $[X] =$  the empty word,  $[2X] = 0$ ,  $[2X + 1] = 1$ ,  $[16X + 3] = 0011$ .

An obvious remark is that  $[p^i X + j] = [p^k X + \ell]$  imply that  $i = k$  and  $j = \ell$ . Besides, one has

$$[(p^i X + j)^{t_s}] = s[p^i X + j] = \text{the word } [p^i X + j] \text{ with an } s \text{ on the left.}$$

As a result, if  $[p^i X + j] = d_{i-1} \dots d_1 d_0$ , then by putting  $w = t_{d_0} t_{d_1} \dots t_{d_{i-1}}$ , we have  $X^w = p^i X + j$ . (Incidentally this shows that  $w$  can be recovered from  $X^w$ , and so  $X^w = X^{w'}$  imply  $w = w'$ . We will not have much use for this remark.)

What is important is that this establishes the claim. The relationship with the connectedness of  $N(a)$  is clear, since we have given a definition of  $u^w$  for  $u \in N(a)$ , and we have  $u^w = u^{(i,j)}$  where  $X^w = p^i X + j$  (the notation  $u^{(i,j)}$  is explained in the introduction). We will sometimes go back and forth between the  $w$ - and  $(i, j)$ -notation.

The operation  $t_1$  is (slightly) more important than the others (since the theory is richer for 2-automatic sequences), so one final comment will be handy. When we put  $X = 1$ , the number  $p^i + j$  is written in base  $p$  as  $[(p^i X + 1)^{t_1}] = 1[p^i X + j] =$  the word  $[p^i X + j]$  with a 1 on the left.

**2.5. The type of a relation; global relations.** The *type* of a relation  $(i, j) \in \text{rel}(a)$  is the leftmost digit in  $[p^i X + j]$ , or equivalently, is 0 when  $j < p^{i-1}$  and is the leftmost digit of  $j$  written in base  $p$  otherwise. Let us say that  $(i, j)$  is a *global relation* for  $a$  if

$$(i, j) \in \bigcap_{u \in N(a)} \text{rel}(u);$$

and let us say that  $a$  has *global relations of all types* if there is (at least) a global relation  $(i, j)$  of type  $r$  for each  $0 \leq r < p$ .

LEMMA 2.3 – *Let  $(a)$  be a  $p$ -automatic sequence.*

- (1) *Let  $r$  be an integer with  $0 \leq r < p$ . Then there is  $u \in N(a)$  such that  $\text{rel}(u)$  contains a relation of type  $r$ .*
- (2) *If  $(a)$  is homogeneous, then  $(a)$  has global relations of all types.*

Property (2) of this lemma will imply that, in practice, we will almost never have to worry about the types of relations.

*Proof.* (1) The elements  $a, a^{tr}, (a^{tr})^{tr} = a^{t_r^2}, \dots, a^{t_r^k}, \dots$ , defined for all  $k \geq 0$ , cannot be all different, since they are taken from the finite set  $N(a)$ . So there must exist  $k \geq 0$  and  $m > 0$  such that

$$a^{t_r^{k+m}} = a^{t_r^k}.$$

In other words  $u^{t_r^m} = u$  where  $u = a^{t_r^k}$ . Let  $(i, j)$  be such that  $X^{t_r^m} = p^i X + j$ , as above ; since  $[p^i X + j] = rrrr \dots$  ( $m$  times), the type of  $(i, j)$  is certainly  $r$ . Moreover  $u^{(i,j)} = u^{t_r^m} = u$ , so  $(i, j) \in \text{rel}(u)$ .

(2) When  $a$  is homogeneous,  $\text{rel}(u)$  is the same for all  $u \in N(a)$ . By the first point, this set contains relations of all types.  $\square$

### 3. THE MAIN THEOREMS

We now have all the tools to embark on a proof of the theorems stated in the introduction. These will be obtained in a completely different order.

**3.1. Sequences with global relations of all types.** These sequences seem perhaps less interesting for their own sake than homogeneous or self-similar sequences (and condition (R1) below is admittedly a bit artificial). However it is technically quite easy to start with their properties, and subsequently refine the results to deal with other types of sequences.

PROPOSITION 3.1 – *Let  $(a)$  be a 2-automatic sequence. Then the following statements are equivalent.*

- (R1) *The sequence  $(a)$  has global relations of all types.*
- (R2) *The monoid  $G(a)$  is a group.*
- (R3) *The sequence  $(a)$  can be produced by an automaton whose underlying graph is the Schreier graph of a group  $G$  with distinguished generators, acting faithfully and transitively on a set  $Q$ .*

Moreover, we have the following minimality statement. Suppose that  $G$ ,  $Q$  and  $\tau$  (the labeling map  $G \rightarrow \Delta$ ) are as in (R3). Then  $|Q| \geq |N(a)|$ , with equality precisely when the condition below holds:

$$(\dagger) \quad \text{if } \tau(q_0^{hg}) = \tau(q_0^g) \text{ for all } g \in G \text{ then } q_0^h = q_0.$$

(Here  $q_0$  is the initial state, and  $q_0^g$  is the image of  $q_0$  under the action of  $g \in G$ .)

In this case  $G$  can be identified with  $G(a)$ , and  $Q$  with  $N(a)$ . This identification preserves the distinguished generators.

Recall that an action of  $G$  is called *faithful* if no element of  $G$  except for the identity acts as the identity permutation. The study of any group action reduces to the study of a faithful action.

*Proof.* Suppose (R1) holds. In graph-theoretic words, condition (R1) says that each vertex of  $N(a)$  is at the end of an arrow bearing the label  $r$ , for each  $r$ . Thus  $t_r$  is a surjective map  $N(a) \rightarrow N(a)$ , and it must be a bijection since  $N(a)$  is finite,  $a$  being automatic. So  $G(a)$  is a group, which is (R2).

Now assume (R2), and let us prove that (R3) holds with  $G = G(a)$  and  $Q = N(a)$ . The Schreier graph  $\Gamma(a)$  can be turned into an automaton if we define  $\tau: N(a) \rightarrow \Delta$  by  $\tau(u) =$  the first term of  $u^{t_1^{-1}}$ . (Indeed the sequence  $u$  is of the form  $u_n = v_{2n+1}$  for a unique  $v \in N(a)$ , or in other words there is an arrow with the label 1 from  $v$  to  $u$ , so surely we must define the label at  $u$  to be  $v_1$ .) We must prove that  $\Gamma(a)$  then produces  $a$ .

Indeed, suppose  $n = d_i d_{i-1} \cdots d_1 d_0$  in base 2, let  $w = t_{d_0} \cdots t_{d_i}$ . Following the arrows accordingly in  $\Gamma(a)$  starting from  $a$  leads us to  $a^w$ . Now we know that  $d_i = 1$ , since  $p = 2$ , so let  $w = w' t_1$  and let  $v = a^{w'}$ , so that  $a^w = v^{t_1}$ . The label  $\tau(a^w)$  is  $v_1$  by definition, and we must show that it is  $a_n$ . However  $v = a^{w'} = a^{(i,j)}$  where  $X^{w'} = 2^i X + j$ , with notation as in §2.4 ; that is  $v_n = a_{2^i n + j}$  for  $n \geq 1$ , so in particular  $v_1 = a_{2^i + j}$ . Here the very last remark in §2.4 is that  $2^i + j$ , when written in base 2, is  $1[2^i X + j] = 1d_{i-1}d_{i-2} \cdots d_1 d_0$ , which is also how  $n$  is written, so  $2^i + j = n$ . This proves (R3).

We check the minimality condition. That  $\tau(a^{hg}) = \tau(a^g)$  for all  $g \in G$  is equivalent to claiming that the sequence obtained from the automaton with  $a^h$  as the initial state is again  $a$ . However, the map from Lemma 2.1 is now a surjective map  $N(a) \rightarrow N(a)$ , which must also be injective, showing that  $a^h = a$ .

That (R3) implies (R1) is almost obvious. The sequence  $a$  is produced, when R3 holds, by an automaton in which each state is at the end of an arrow marked  $r$ , for all  $r$ . Using the surjective map  $Q \rightarrow N(a)$  from Lemma 2.1, we deduce that any  $u \in N(a)$  is of the form  $v^{t_r}$ , for all  $r$ , and that is condition (R1) reworded.

Here we use the fact that the map  $Q \rightarrow N(a)$  is compatible with the generators  $t_i$ , in the sense that if it maps  $q$  to  $u$ , then it maps  $q^{t_i}$  to  $u^{t_i}$  ; this holds obviously, even when there is no group in sight and  $q^{t_i}$  is simply taken to mean the state at the end of the arrow with label  $i$  originating at  $q$ .

The same observation will prove the last statements. For suppose condition  $(\dagger)$  is satisfied. We claim that the map  $Q \rightarrow N(a)$  is then injective. Since the action of  $G$  is transitive, any  $q \in Q$  is of the form  $q_0^h$ , where  $q_0$  is the initial state. Suppose  $h, k \in G$  are such that the sequences produced by choosing  $q_0^h$  and  $q_0^k$  as initial state, respectively, coincide. Then  $\tau(q_0^{hx}) = \tau(q_0^{kx})$  for all  $x \in G$ , and in particular for  $x = k^{-1}g$  we find  $\tau(q_0^{h^{k^{-1}g}}) = \tau(q_0^g)$  for all  $g \in G$ , so  $q_0^{h^{k^{-1}}} = q_0$  by  $(\dagger)$ , and  $q_0^h = q_0^k$ . This proves the claim.

This provides the required identification of  $Q$  with  $N(a)$ , compatible with the bijections denoted by  $t_i$  on each of these sets. The corresponding permutation groups can also be identified, and these are  $G$  (because the action is assumed to be faithful) and  $G(a)$  (by definition).

It remains to note that when  $(\dagger)$  does not hold, the map  $Q \rightarrow N(a)$  is not injective (but still surjective), so  $Q$  has more elements than  $N(a)$ .  $\square$

Large parts of this proof work in the general case of  $p$ -automatic sequences for any  $p$ . We state this separately:

**PROPOSITION 3.2** – *Let  $G$  be a finite group generated by  $t_0, \dots, t_{p-1}$ , and suppose that  $G$  acts faithfully and transitively on a finite set  $Q$ . Suppose that the corresponding Schreier graph is turned into an automaton by choosing an initial state  $q_0 \in Q$  and a labeling map  $\tau: Q \rightarrow \Delta$  satisfying the same minimality condition  $(\dagger)$  as above. Let  $(a)$  be the  $p$ -automatic sequence produced.*

*Then  $(a)$  has global relations of all types, the monoid  $G(a)$  is a group which can be identified with  $G$ , and  $\Gamma(a)$  can be identified with the graph underlying the automaton.*

**EXAMPLE 3.3** – We prove that Proposition 3.1 does not hold for  $p > 2$ . Indeed for any  $p$ , let  $a_n =$  the last digit of  $n$  written in base  $p$ . All the subsequences of

the form  $(a_{p^i n + j})_{n \geq 1}$  coincide with the original sequence, so  $N(a) = \{a\}$  and  $a$  is  $p$ -automatic. However the graph  $\Gamma(a)$  has only one vertex, and no matter the decoration we elect to place on  $\Gamma(a)$ , the automaton obtained can only produce the constant sequence. This will be different from  $a$  except for  $p = 2$ . So  $a$  cannot be produced by  $\Gamma(a)$ , even though it certainly has all the relations one could ask for (so R1 holds).

### 3.2. Homogeneous sequences.

**THEOREM 3.4** – *Let  $(a)$  be a 2-automatic sequence. Then the following statements are equivalent.*

- (H1) *The sequence  $(a)$  is homogeneous.*
- (H2) *The monoid  $G(a)$  is a group, and  $\Gamma(a)$  is its Cayley graph.*
- (H3) *The sequence  $(a)$  can be produced by an automaton whose underlying graph is the Cayley graph of a group  $G$  with distinguished generators (with 1 as the initial state), and with the property that the subgroup  $H$  of elements satisfying  $\tau(hg) = \tau(g)$  for all  $g \in G$  is normal in  $G$ .*

*Moreover in (H3) we can arrange to have  $G = G(a)$ , and  $H = \{1\}$ . Conversely, suppose we start with  $G$  as in (H3) and that  $H = \{1\}$ . Then  $G$  can be identified with  $G(a)$ . This identification preserves the distinguished generators.*

*Proof.* We start by noting that (H1) implies (R1) (Lemma 2.3), (H2) implies (R2) and (H3) implies (R3) (trivially). So by Proposition 3.1, we may conduct the proof assuming that all three properties (R1-R2-R3) hold.

Assume (H1). For  $u \in N(a)$ , we let  $\text{Stab}(u)$  denote its stabilizer under the action of  $G(a)$  (that is, the group of  $g \in G(a)$  such that  $u^g = u$ ). The fact that  $\Gamma(a)$  is connected (or equivalently, that the action of  $G(a)$  is transitive) implies that the various groups  $\text{Stab}(u)$  are all conjugate as  $u$  runs through  $N(a)$ . Also, the intersection of all these is the trivial subgroup  $\{1\}$ , since it consists of elements  $g$  fixing everything in  $N(a)$ , while  $G(a)$  is a group of permutations of this set by definition.

However (H1) says that  $\text{rel}(u)$  does not depend on the choice of  $u$ , and it follows that  $\text{Stab}(u)$  is also independent of the particular subsequence  $u$ . Finally we see that the subgroups  $\text{Stab}(u)$ , being all equal with trivial intersection, are all trivial. For cardinality reasons, this implies that  $|N(a)| = |G(a)|$ , and so  $\Gamma(a)$  can be identified with the Cayley graph of  $G(a)$  (Lemma 2.2). Thus (H2) holds.

When (H2) is assumed, Proposition 3.1 gives (H3) with  $G = G(a)$  and  $H = \{1\}$  (so that, in particular,  $H$  is normal in  $G$ ). The “moreover” statements will also be clear, but we should finish the equivalence first.

So assume (H3). Replacing  $G$  by  $G/H$  if necessary (which makes sense since  $H$  is assumed to be normal), we are reduced to the case  $H = \{1\}$ . By Proposition 3.1, we can and we do identify  $G$  with  $G(a)$  and  $Q$  with  $N(a)$ . Following arrows leading from any  $u \in N(a)$  to  $a$  (and in a Cayley graph, this is always possible), we see that  $a \in N(u)$ , so  $N(a) = N(u)$ . Further, consider the condition  $(i, j) \in \text{rel}(u)$ . It is equivalent to the requirement that the element of  $G(a)$  obtained as the word in the generators  $t_r$  corresponding to the word  $[2^i X + j]$  as in §2.4 be an element of  $\text{Stab}(u)$  (this is just a game with notation). However for all  $u$  we have  $\text{Stab}(u) = \{1\}$  since we are in a Cayley graph, and we see that the condition  $(i, j) \in \text{rel}(u)$  actually does not depend on  $u$ . So  $\text{rel}(u) = \text{rel}(a)$ . We have (H1).  $\square$

For general  $p$ -sequences, what remains true is:

**THEOREM 3.5** – *Let  $G$  be a group with distinguished generators  $t_0, \dots, t_{p-1}$ . Suppose the corresponding Cayley graph is turned into an automaton, with initial state 1, by means of a map  $\tau: G \rightarrow \Delta$  such that the subgroup  $H$  of elements*

satisfying  $\tau(hg) = \tau(g)$  for all  $g \in G$  is normal in  $G$ . Let  $(a)$  be the  $p$ -automatic sequence produced.

Then  $a$  is homogeneous, the monoid  $G(a)$  is a group and  $\Gamma(a)$  is its Cayley graph. If moreover  $H = \{1\}$ , then  $G$  can be identified with  $G(a)$ , and  $\Gamma(a)$  can be identified with the graph underlying the automaton.

### 3.3. Self-similar sequences.

**THEOREM 3.6** – *Let  $a$  be a 2-automatic sequence, on an alphabet  $\Delta$  (we assume that all the letters of  $\Delta$  are actually used). The following statements are equivalent.*

- (S1) *The sequence  $(a)$  is self-similar.*
- (S2) *The sequence  $(a)$  can be obtained by an automaton whose underlying graph is the Cayley graph of a group  $G$  with distinguished generators (with 1 as the initial state), and whose labeling map is the natural map  $\tau: G \rightarrow G/K \cong \Delta$  for some subgroup  $K$ .*

Moreover, if the intersection of all the conjugates of  $K$  is trivial, then  $G$  can be identified with  $G(a)$ .

We stress that  $G/K$  is not  $K \backslash G$ : it is the set of classes  $gK$  for  $g \in G$ , and there is an action *on the left* on this set. Other actions so far have been *on the right*. This seems inevitable.

*Proof.* Assume (S1). The sequence  $(a)$ , being self-similar, is also homogeneous. From Theorem 3.4, it can be produced by the automaton with underlying graph  $\Gamma(a)$ , the Cayley graph of the group  $G = G(a)$ , and we view  $\tau$  as defined on  $G$ . Now any  $u \in N(a)$  is of the form  $a^h$  for a *unique*  $h \in G$ . The sequence  $a$  is given by  $a_n = \tau(g_n)$  where  $g_n \in G$  is written as the word in the  $t_r$ 's corresponding to the digits of  $n$  is base 2; the sequence  $a^h$  is given by  $a_n^h = \tau(hg_n)$ .

By hypothesis, to each  $h$  we can attach a bijection  $\varphi_h: \Delta \rightarrow \Delta$  such that  $\varphi_h(a_n) = a_n^h$  for all  $n \geq 1$ , or  $\tau(hg_n) = \varphi_h(\tau(g_n))$ . Let us write  $h \cdot \delta$  instead of  $\varphi_h(\delta)$ , for  $\delta \in \Delta$ , so that

$$\tau(hg) = h \cdot \tau(g)$$

for all  $g \in G$ . The map  $\tau$  was assumed to be surjective, so this last relation ensures that  $h, \delta \rightarrow h \cdot \delta$  is a (left) action of  $G$  on  $\Delta$ ; moreover this action is compatible with  $\tau$  and the left action of  $G$  on itself. The latter action is transitive, and thus, so must be the former. If  $K$  is the stabilizer of any point in  $\Delta$ , we can then identify  $\Delta$  with  $G/K$ . We have proved (S2).

The proof that (S2) implies (S1) is trivial: we *define*  $\varphi_h$  to be the action of  $h$  on  $\Delta = G/K$ .

The “moreover” statement follows from that of Theorem 3.4. □

**THEOREM 3.7** – *Let  $G$  be a group with distinguished generators  $t_0, \dots, t_{p-1}$ . Suppose the corresponding Cayley graph is turned into an automaton, with initial state 1, by means of the map  $\tau: G \rightarrow G/K$  for some subgroup  $K$ . Let  $(a)$  be the  $p$ -automatic sequence produced.*

*Then  $(a)$  is self-similar, the monoid  $G(a)$  is a group and  $\Gamma(a)$  is its Cayley graph. If moreover the intersection of all the conjugates of  $K$  is trivial, then  $G$  can be identified with  $G(a)$ , and  $\Gamma(a)$  can be identified with the graph underlying the automaton.*

## 4. RATIONAL FRACTIONS ASSOCIATED TO AUTOMATIC SEQUENCES

**4.1. Defining the rational fractions.** Suppose  $(a_n)_{n \geq 1}$  is a  $p$ -automatic sequence on an alphabet which is a subset of a ring  $R$ . We define

$$L(a, x_0, \dots, x_{p-1}) = \sum_{n \geq 1} a_n x_0^{\ell_0(n)} \cdots x_{p-1}^{\ell_{p-1}(n)} \in R[[x_0, \dots, x_{p-1}]]$$

where  $\ell_i(n)$  is the number of occurrences of the digit  $i$  when writing  $n$  in base  $p$ . We also define

$$L(a, x) = L(a, x, x, \dots, x) = \sum_{n \geq 1} a_n x^{\ell(n)} \in R[[x]]$$

where  $\ell(n)$  is the length of  $n$  when written in base  $p$ .

**THEOREM 4.1** – *Under the assumption that  $a$  is  $p$ -automatic, the power series  $L(a, x_0, \dots, x_{p-1})$  is a rational fraction. In fact one can write  $L(a, x_0, \dots, x_{p-1})$  as a coefficient of the column vector*

$$(I - M)^{-1}C$$

where  $M \in \mathbb{Z}[x_0, \dots, x_{p-1}]$  is the weighted adjacency matrix of the graph  $\Gamma(a)$ , we write  $I$  for the identity matrix, and  $C$  is a column vector whose entries are homogeneous polynomials in  $R[x_1, \dots, x_{p-1}]$  of degree 1 (that is, linear forms).

During the course of the proof we shall make precise the term “weighted adjacency matrix”, as well as give an expression for  $C$ .

*Proof.* Let  $u \in N(a)$ . We write  $L = L(u, x_0, \dots, x_{p-1})$  and  $L^{t_i} = L(u^{t_i}, x_0, \dots, x_{p-1})$ . Note that

$$\sum_{n \geq 1} u_{pn+i} x_0^{\ell_0(pn+i)} \cdots x_{p-1}^{\ell_{p-1}(pn+i)} = x_i L^{t_i}.$$

As a result, by partitioning the integers according to their values mod  $p$ , one has

$$(*) \quad L = u_1 x_1 + \cdots + u_{p-1} x_{p-1} + x_1 L^{t_1} + \cdots + x_{p-1} L^{t_{p-1}}.$$

Now fix an arbitrary order on  $N(a)$ . We shall work with matrices and vectors which are indexed by the elements of  $N(a)$ , with this order.

We start with the matrix  $M$  defined by

$$M_{u,v} = \sum_{u^{t_i}=v} x_i,$$

which we call “the weighted adjacency matrix of  $\Gamma(a)$ ” (see also the definition of the matrix  $A$  below). Next, define the column vectors  $\Lambda$  and  $C$  by

$$\Lambda_u = L(u, x_0, \dots, x_{p-1}), \quad C_u = u_1 x_1 + \cdots + u_{p-1} x_{p-1}.$$

Equation  $(*)$  above can now be written

$$(I - M)\Lambda = C.$$

It remains to prove that  $I - M$ , a matrix with entries in  $\mathbb{Z}[x_0, \dots, x_{p-1}]$ , is invertible in the field of fractions  $\mathbb{Q}[x_0, \dots, x_{p-1}]$ . For this it suffices to check that its determinant is not the zero polynomial, and in turn, it suffices to show this after evaluating at  $x_0 = x_1 = \cdots = x_{p-1} = x$ .

In this situation  $M = xA$  where  $A_{u,v}$  is the number of indices  $i$  such that  $u^{t_i} = v$ , and we call  $A$  the adjacency matrix of  $\Gamma(a)$ . Now

$$\det(I - xA) = x^N \det\left(\frac{1}{x}I - A\right) = x^N \chi_A\left(\frac{1}{x}\right).$$

Here  $N = |N(a)|$  and  $\chi_A$  is the characteristic polynomial of  $A$ , which has degree  $N$ . The proof is complete.  $\square$

COROLLARY 4.2 – The power series  $L(a, x)$  is a rational fraction, given by a coefficient of the column vector

$$(I - xA)^{-1}xT,$$

where  $A$  is the adjacency matrix of  $\Gamma(a)$  as defined above, and  $T_u = u_1 + \dots + u_{p-1}$ .

**4.2. Examples.** We shall examine our two running examples. From now on when we deal with a sequence on an alphabet  $\{d_1, \dots, d_r\}$ , we assume that the ring  $R$  is  $\mathbb{Z}[d_1, \dots, d_r]$ .

EXAMPLE 4.3 – The Thue-Morse sequence  $a$  is produced by the automaton on Figure 1. It is a Cayley graph, for the group of order two  $G = \{\pm 1\}$  with  $t_0 = 1$  and  $t_1 = -1$ . The labeling map  $\tau$  is a bijection between the elements of  $G$  and  $\{A, B\}$ . Theorem 3.7 implies that  $a$  is self-similar and, most importantly here,  $G(a)$  is nothing but  $G$  itself, and  $\Gamma(a)$  is the directed graph underlying the automaton. There are 2 sequences in  $N(a)$ , corresponding to the two possible initial states, and these are  $a$  and  $a^{t_1} = a^{(1,1)} = (a_{2n+1})_{n \geq 1}$ . Say  $N(a)$  is ordered so that  $a$  is the first element.

The matrix  $M$  is thus, using the letters  $x$  and  $y$  rather than  $x_0$  and  $x_1$ :

$$M = \begin{pmatrix} x & y \\ y & x \end{pmatrix}.$$

Writing  $L$  and  $L^{t_1}$  for the rational fractions associated to  $a$  and  $a^{t_1}$  respectively, we have

$$\Lambda = \begin{pmatrix} L \\ L^{t_1} \end{pmatrix}, \quad C = \begin{pmatrix} xa_1 \\ xa_1^{t_1} \end{pmatrix} = \begin{pmatrix} xB \\ xA \end{pmatrix}.$$

We only need to compute

$$(I - M)^{-1} = \frac{1}{-x^2 + y^2 + 2x - 1} \begin{pmatrix} x - 1 & -y \\ -y & x - 1 \end{pmatrix}$$

so that

$$\Lambda = (I - M)^{-1}C = \frac{1}{-x^2 + y^2 + 2x - 1} \begin{pmatrix} -xyA + (x^2 - x)B \\ (x^2 - x)A - xyB \end{pmatrix}$$

and in particular

$$L(a, x, y) = \frac{-xyA + (x^2 - x)B}{-x^2 + y^2 + 2x - 1}$$

and

$$L(a, x) = L(a, x, x) = \frac{-x^2A + (x^2 - x)B}{2x - 1}.$$

Notice how  $L(a^{t_1}, x)$  is obtained from  $L(a, x)$  by exchanging  $A$  and  $B$ , and likewise for  $L(a^{t_1}, x, y)$  and  $L(a, x, y)$ .

EXAMPLE 4.4 – Let  $b$  be the 7-automatic sequence produced by the automaton on Figure 2. We compute  $L(b, x)$  (from now on we shall be interested in single-variable rational fractions, for simplicity).

We have already observed in the introduction that the underlying graph is the Cayley group for a group  $G$ , which is cyclic of order 6, and we have specified the generators  $t_0, \dots, t_6$  as permutations. The 6 labels are all distinct, so we have a self-similar sequence, as dealt with by Theorem 3.7. The set  $N(a)$  is in bijection with the set of states, and since these are labeled  $d_1, \dots, d_6$ , we have a natural

order. The adjacency matrix is then

$$A = \begin{pmatrix} 2 & 0 & 3 & 0 & 2 & 0 \\ 0 & 2 & 2 & 0 & 0 & 3 \\ 2 & 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 & 2 \\ 3 & 0 & 0 & 2 & 2 & 0 \\ 0 & 2 & 0 & 3 & 0 & 2 \end{pmatrix}.$$

To construct the vector  $T$ , for each integer  $i$  we compute the indices  $j_1 = t_1(i), \dots, j_6 = t_6(i)$ , and the  $i$ -th line of  $T$  is  $d_{j_1} + \dots + d_{j_6}$ . Explicitly

$$T = \begin{pmatrix} d_1 + 3d_3 + 2d_5 \\ d_2 + 2d_3 + 3d_6 \\ 2d_1 + 3d_2 + d_3 \\ d_4 + 3d_5 + 2d_6 \\ 3d_1 + 2d_4 + d_5 \\ 2d_2 + 3d_4 + d_6 \end{pmatrix}.$$

The fraction  $L(b, x)$  is then the first line of  $(I - xA)^{-1}xT$ . This straightforward task was performed by a computer, and the result is that which we have given in the introduction.

**4.3. Applications to the frequency of letters.** The fraction  $L(a, x)$  is useful in providing information on the frequency of letters within  $a$ . Recall that we write  $a[n, d_i]$  for the number of occurrences of  $d_i$  among the first  $n$  terms of  $a$ . The next lemma provides the connection, and its proof is immediate.

LEMMA 4.5 – *The expansion of  $L(a, x)$  is*

$$L(a, x) = \sum_{n \geq 1} (m_{1,n}d_1 + \dots + m_{r,n}d_r)x^n$$

where  $m_{i,n}$  is the number of integers  $m$ , having length  $n$  when written in base  $p$ , such that  $a_m = d_i$ . As a result, if

$$L(a, x) = \frac{d_1P_1 + \dots + d_rP_r}{D}$$

with  $P_i, D \in \mathbb{Z}[x]$ , and if we write

$$\frac{P_i}{D} = \sum_{n \geq 1} s_n x^n,$$

then

$$\sum_{j=1}^n s_j = a[p^n, d_i]. \quad \square$$

We now make a few technical points.

PROPOSITION 4.6 – *Let  $D \in \mathbb{C}[x]$ . Assume that the roots of  $D$  have absolute value  $\geq \frac{1}{p}$ , and let the roots of absolute value  $\frac{1}{p}$  be written  $\alpha_k = \frac{1}{p}e^{i\theta_k}$  for  $k = 1, 2, \dots$ . Assume further that each  $\alpha_k$  is a simple root, and that  $D(0) \neq 0$ . Finally, let  $N \in \mathbb{C}[x]$  be any polynomial.*

*If we write*

$$\frac{N}{D} = \sum_{n \geq 0} s_n x^n,$$

*then we have the following estimate:*

$$\frac{\sum_{j=1}^n s_j}{p^n} = \sum_k \frac{N(\alpha_k) \operatorname{Res}(\frac{1}{D}, \alpha_k)}{\alpha_k^2 - \alpha_k} e^{-ni\theta_k} + o(1).$$



*Proof.* We examine a few particular cases for  $D$  and  $N$ . Suppose first that  $N = 1$  and

$$\frac{1}{D} = \frac{1}{(x - \alpha)^m} = \frac{c}{(1 - \gamma x)^m}$$

where  $\gamma = \frac{1}{\alpha}$  and  $c = \frac{1}{\alpha^m}$  is constant. Here we assume that  $|\alpha| > \frac{1}{p}$  so  $|\gamma| < p$ . In this case  $s_n = P(n)\gamma^n$  where  $P$  is a polynomial of degree  $< m$ . Choose a constant  $C > 0$  such that  $|P(n)| \leq Cn^{m-1}$  for all  $n \geq 1$ . Thus

$$\left| \frac{\sum_{j=1}^n s_j}{p^n} \right| \leq \frac{Cn^m |\gamma| (1 + |\gamma| + \dots + |\gamma|^{n-1})}{p^n} = \frac{C|\gamma|n^m}{|\gamma| - 1} \cdot \frac{|\gamma|^n - 1}{p^n} = o(1).$$

Now suppose  $\alpha = \frac{1}{p}e^{i\theta}$ , that  $N = 1$  and that

$$\frac{1}{D} = \frac{1}{X - \alpha} = -\frac{\gamma}{1 - \gamma x}$$

where  $\gamma = \frac{1}{\alpha}$ , so that  $s_n = -\gamma^{n+1}$ . Now

$$\begin{aligned} \frac{\sum_{j=1}^n s_j}{p^n} &= \frac{\gamma^2}{\gamma - 1} \cdot \frac{\gamma^n - 1}{p^n} \\ &= \frac{\gamma^2}{1 - \gamma} (e^{-ni\theta} + o(1)) \\ &= \frac{1}{\alpha^2 - \alpha} e^{-ni\theta} + o(1). \end{aligned}$$

The general case is obtained by writing the partial fraction decomposition of  $\frac{N}{D}$ .  $\square$

The Proposition and the Lemma together show:

**THEOREM 4.7** – Suppose that  $a$  is  $p$ -automatic, on the alphabet  $\{d_1, \dots, d_r\}$ , and that

$$L(a, x) = \frac{d_1 P_1 + \dots + d_r P_r}{D}$$

with  $P_i, D \in \mathbb{Z}[x]$ . Assume that the (complex) roots of  $D$  have absolute value  $\geq \frac{1}{p}$ , and let the roots of absolute value  $\frac{1}{p}$  be written  $\alpha_k = \frac{1}{p}e^{i\theta_k}$  for  $k = 1, 2, \dots$ . Assume further that each  $\alpha_k$  is a simple root.

Then one has

$$\frac{a[p^n, d_i]}{p^n} = \sum_k \frac{P_i(\frac{1}{p}) \text{Res}(\frac{1}{D}, \alpha_k)}{\alpha_k^2 - \alpha_k} e^{-ni\theta_k} + o(1).$$

**COROLLARY 4.8** – Suppose that the only root of  $D$  of absolute value  $\frac{1}{p}$  is  $\frac{1}{p}$  (and that this root is simple). Then

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{a[p^n, d_i]}{p^n} &= \frac{p^2}{1 - p} P_i\left(\frac{1}{p}\right) \text{Res}\left(\frac{1}{D}, \frac{1}{p}\right) \\ &= \frac{P_i(\frac{1}{p})}{\sum_j P_j(\frac{1}{p})}. \end{aligned}$$

*Proof.* The first equality follows directly from the Theorem, while the second is drawn by observing that the various limits, as  $i$  ranges from 1 to  $r$ , all exist and sum up to 1.  $\square$

**EXAMPLE 4.9** – As explained in the introduction, the corollary applies to the sequence  $b$  of Apéry numbers mod 7, produced by the automaton on Figure 2. The polynomials  $P_i$ , different as they are, give the same value at  $\frac{1}{7}$ . It follows that

$$\lim_{n \rightarrow +\infty} \frac{b[7^n, d_i]}{7^n} = \frac{1}{6}.$$

4.4. **A complete example.** Consider the 2-automaton below.

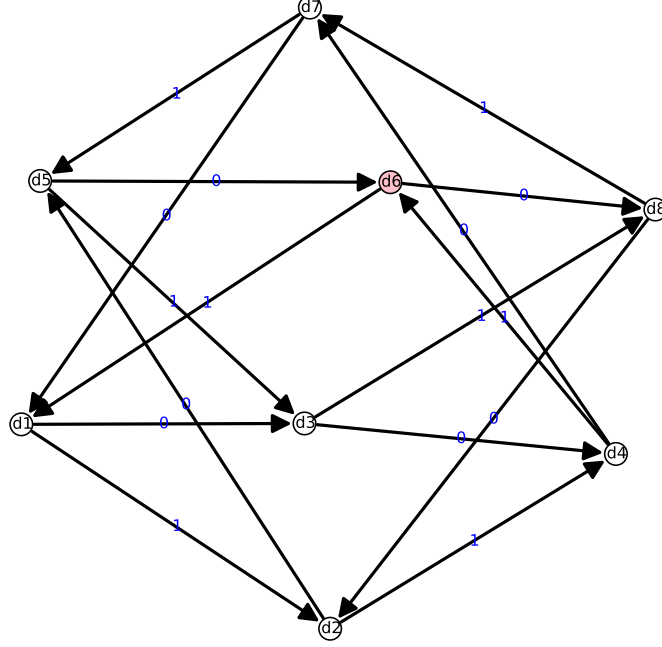


Fig. 3

We shall study the corresponding sequence  $a$ , and indulge in all the details.

The first thing to notice is that each state has two incoming arrows with labels 0 and 1, and also two outgoing arrows with labels 0 and 1, so we are in the presence of a Schreier graph. To check that it is in fact a Cayley graph, we have no choice but consider the group  $G$  generated by the permutations

$$t_0 = (1, 3, 4, 7)(5, 6, 8, 2) \quad \text{and} \quad t_1 = (1, 2, 4, 6)(5, 3, 8, 7).$$

One checks that  $G$  has order 8, and indeed is isomorphic to the group of quaternions (for example by asking a computer). Lemma 2.2 guarantees that we have a Cayley graph. The labeling map  $\tau: G \rightarrow \{d_1, \dots, d_8\}$  is injective, so the alphabet is essentially  $G$  itself, and we have a self-similar sequence. By Theorem 3.7, the graph  $\Gamma(a)$  is simply that on Figure 3.

To proceed with the computations of  $L(a, x)$ , we order the sequences in  $N(a)$  according to the labels on their initial states, so that the first sequence is that with initial state  $d_1$ , the second is that with initial state  $d_2$ , and so on. The sequence  $a$  itself, according to the picture, has initial state  $d_6$ , so it is the sixth in this order.

The adjacency matrix is then

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

and the vector  $T$  is:

$$T = \begin{pmatrix} d_2 \\ d_4 \\ d_8 \\ d_6 \\ d_3 \\ d_1 \\ d_5 \\ d_7 \end{pmatrix}.$$

(because the first sequence starts with  $d_2$ , the second with  $d_4$ , etc). The fraction  $L(a, x)$  is the 6-th line in the vector

$$(I - xA)^{-1}xT,$$

and we find

$$L(a, x) = \frac{d_1P_1 + \dots + d_8P_8}{8x^4 + 2x^2 - 1}$$

with

$$\begin{aligned} P_1 &= 2x^5 + 2x^3 - x, & P_2 &= x^4 - x^2, \\ P_3 &= -3x^4, & P_4 &= 2x^5 - 2x^3, \\ P_5 &= -2x^5 - x^3, & P_6 &= -3x^4, \\ P_7 &= x^4 - x^2, & P_8 &= -2x^5 - x^3. \end{aligned}$$

So the denominator is

$$D = (2x - 1) \cdot (2x + 1) \cdot (2x^2 + 1),$$

its roots being  $\frac{1}{2}$ ,  $-\frac{1}{2}$ ,  $i\frac{\sqrt{2}}{2}$  and  $-i\frac{\sqrt{2}}{2}$ , the last two having modulus  $\frac{\sqrt{2}}{2} > \frac{1}{2}$ . The residue of  $\frac{1}{D}$  at  $\pm\frac{1}{2}$  is  $\pm\frac{1}{6}$ .

Thus we have

$$\frac{a[2^n, d_i]}{2^n} = -\left[\frac{2}{3}P_i\left(\frac{1}{2}\right) + (-1)^n\frac{2}{9}P_i\left(-\frac{1}{2}\right)\right] + o(1).$$

For  $i \in \{1, 4, 5, 8\}$ , we have  $P_i(\frac{1}{2}) = -\frac{3}{16}$  and  $P_i(-\frac{1}{2}) = \frac{3}{16}$  so

$$\frac{a[2^n, d_i]}{2^n} = \frac{1}{8} + (-1)^{n+1}\frac{1}{24} + o(1).$$

For these indices, we conclude that when  $n$  is large and even, the ratio  $a[2^n, d_i]/2^n$  is close to  $\frac{1}{8} - \frac{1}{24} = \frac{1}{12}$ , but when  $n$  is odd the ratio is close to  $\frac{1}{8} + \frac{1}{24} = \frac{1}{6}$ .

For  $i \in \{2, 3, 6, 7\}$ , on the other hand, we find  $P_i(\frac{1}{2}) = P_i(-\frac{1}{2}) = -\frac{3}{16}$ , and so

$$\frac{a[2^n, d_i]}{2^n} = \frac{1}{8} + (-1)^n\frac{1}{24} + o(1).$$

The ratio in this case is close to  $\frac{1}{6}$  when  $n$  is even and large, and close to  $\frac{1}{12}$  if  $n$  is odd and large.

In particular, none of the ratios  $\frac{a[2^n, d_i]}{2^n}$  converges. Also note that when  $n$  is even, approximately two thirds of the first  $2^n$  terms of the sequence  $a$  are in the set  $\{d_2, d_3, d_6, d_8\}$ . When  $n$  is odd, exactly the opposite is true.

## REFERENCES

- [AS03] Jean-Paul Allouche and Jeffrey Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003, Theory, applications, generalizations. MR 1997038 (2004k:11028)
- [Gui14] Pierre Guillot, *An elementary approach to dessins d'enfants and the Grothendieck-Teichmüller group*, L'enseignement Mathématique **60** (2014), 293–375.
- [Row15] Eric Rowland, *What is ... an automatic sequence?*, Notices Amer. Math. Soc. **62** (2015), no. 3, 274–276. MR 3308393
- [RY15] Eric Rowland and Reem Yassawi, *Automatic congruences for diagonals of rational functions*, J. Théor. Nombres Bordeaux **27** (2015), no. 1, 245–288. MR 3346972

UNIVERSITÉ DE STRASBOURG & CNRS, INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE,  
7 RUE RENÉ DESCARTES, 67084 STRASBOURG, FRANCE  
*E-mail address:* guillot@math.unistra.fr